User Grouping and Load Balancing for FDD Massive MIMO Systems

Yi Xu[1]*, Guosen Yue*[2] *and Shiwen Mao*[1]
[1]*Auburn University, Auburn, AL, USA*
[2]*Broadcom Corporation, Matawan, NJ, USA*
yzx0010@auburn.edu, guosenyue@gmail.com, smao@ieee.org

1. Introduction

Last decades have witnessed ever-increasing demand for higher data rates in wireless networks. To cater for this demand, many advanced physical layer techniques have been developed, e.g., multiple input multiple output (MIMO) with orthogonal frequency division multiplexing (OFDM). However, with linear throughput improvement but the exponential growth on the data traffic, the gap between the demand and supply has been increasingly widened. To solve the problem, the next technology we could resort to is massive MIMO (a.k.a. large-scale MIMO, full-dimension MIMO, or hyper MIMO), which significantly increases the system capacity by employing a large number of antennas at the base station. As an emerging and promising technology, large-scale MIMO also enjoys many advantages such as low-power, robust transmissions, simplified transceiver design, and simplified multiple access layer [1], [2], in addition to enhanced capacity.

In general, the more transmit antennas, the more degrees of freedom a massive MIMO system can provide, resulting in higher reliability or larger throughput. However, due to the difficulties of acquiring channel state information at the transmitter side (CSIT), it is challenging to simultaneously support a large number of users [2]. Most of the existing works on massive MIMO systems consider the time division duplexing (TDD) mode [3]-[5], within which by exploiting channel reciprocity, the downlink channel can be estimated through uplink training. Unfortunately, there is no such privilege in frequency-division-duplexing (FDD) systems.

There are much more FDD (≥ 300) than TDD (≤ 40) LTE licenses worldwide. It is therefore of great importance to investigate the massive MIMO design for FDD systems. To reduce pilot resources and the channel state information (CSI) feedback in FDD systems, a two-stage precoding scheme has been proposed in [6] recently. Firstly, the users in service are divided into groups, while each group of users have similar second-order channel statistics (i.e., transmit correlation). The same pre-beamforming, or the first-stage precoding, is then used for each group of users semi-statically. Next, with reduced dimensions on the effective channel, simplified channel feedback can be realized and the second-stage dynamic precoding can be applied. The performance of such system design is largely dependent on user grouping. In [7], a K-means clustering scheme, based on chordal distance as the clustering metric, is introduced for user grouping. In this letter, instead of chordal distance, we propose weighted likelihood similarity measure and hierarchical clustering. By theoretical analysis and simulations, we validate the proposed approaches.

Once user groups are formed, another important issue is user scheduling, i.e., selecting users for transmission based on instantaneous channel conditions. In this letter, we propose a dynamic user scheduling method. If there are only a few active users, some groups may barely have users while some other groups are overloaded. Therefore, we also consider the load balancing problem and develop an effective solution algorithm. Note that some results can also be found in [8][9].

2. Problem Statement and Main Results

We consider a downlink system with M antennas at the base station (BS) and a single antenna at each user terminal (UT). Denote y_k as the received signal at user$k$, $k = 1,2, \ldots, K$. The signals received by all UTs y can be written as

$$\boldsymbol{HBP}d+z, \quad (1)$$

where $()^H$ denotes the Hermitian of a matrix; $\mathbf{H}$, of dimension $M \times K$, is the actual channel between the BS and the users; $\boldsymbol{V}$ is the precoding matrix of dimension $M \times S$; d is the data vector of dimension $S\times 1$; and z is the zero mean circulant symmetric complex Gaussian noise vector. The key idea is to decompose $\boldsymbol{V}$ into $\boldsymbol{B}$ and $\boldsymbol{P}$, where $\boldsymbol{B}$ is pre-beamforming matrix of dimension $M \times b$; $\boldsymbol{P}$ of dimension $b \times S$, is designed to suppress the interferences within each group.

1) User grouping schemes.

To design the pre-beamforming matrix $\boldsymbol{B}$, we need to group users. Different from the chordal distance based K-means user grouping scheme [7], we first propose a weighted likelihood function as the similarity measure between a user and a group, which is defined as:

$$2 \quad (2)$$

In addition to the new similarity measure, we also propose new user grouping schemes, which employs the agglomerative hierarchical clustering method. Different from the K-means approach, which essentially looks at all possible combinations of users and groups, the agglomerative hierarchical clustering method starts with each individual user forming a user group. It then proceeds by a series of successive mergers based on certain criteria. Eventually, all users can form one single group. We can terminate the scheme when the desired number of groups is reached.

One important issue in hierarchical clustering is how to define the similarity measure or distance between existing groups and newly defined groups (called linkage methods). We propose to use weighted average linkage as follows.

$$q)/2, \quad (3)$$

q are groups.

Given chordal distance, weighted likelihood similarity, K-means clustering and hierarchical clustering, we could combine either two of them to form a complete user grouping scheme.

Proposition 1: The complexity of K-means clustering 2)) for weighted likelihood similarity measure, where r^* is the effective rank of R_k, i.e., the number of k.

Proposition 2: The complexity of hierarchical clustering is $O(K(K-1)(2M^3+M^2)/2)$ for chordal 2)/2) for weighted likelihood similarity measure.

2) User scheduling scheme.

With user groups being formed, we can obtain the pre-beamforming matrix $\boldsymbol{B}_g$ for each group g. At a particular time slot, based on the instantaneous channel conditions of the users, we dynamically schedule a subset of users in each group for the transmissions in this time slot.

In [7], a MAX and an ALL user scheduling algorithm are presented. Different from these approaches, we propose a dynamic user scheduling algorithm that schedules users in a greedy manner. In particular, at each step, the proposed algorithm only schedules the user that can achieve the largest gain in the system throughput. The proposed algorithm is presented in the Algorithm 1.

3) Load balancing scheme.

In practical applications, many users may gather at one geographic location (e.g., in a skyscraper). If we design the precoder exactly as discussed, these users will form a big group. It would be desirable to offload some of the users to other groups, to achieve fairness among the users. This is because with more members in a group, each member's chance of getting scheduled for transmission will be smaller.

```
User groups $\{\mathcal{S}_g\}$ are given ;
Initially set $\mathcal{U} = \{1, 2, \cdots, K\}$, $\mathcal{C} = 0$, and $\mathcal{K}_g = \emptyset, \forall g$ ;
while Termination conditions ($\sum_g |\mathcal{K}_g| = \sum_g b_g$,
$\kappa(k^*, g_{k^*}) = 0$, or $\mathcal{U} = \emptyset$) are not satisfied do
    for $k \in \mathcal{U}$ do
        if $|\mathcal{K}_{g_k}| < S_g$ then
            Set $\mathcal{K}'_g = \mathcal{K}_g \cup \{k\}$ if $k \in \mathcal{S}_g$, and
            $\mathcal{K}'_{g'} = \mathcal{K}_{g'}, \forall g' \neq g$ ;
            Perform ZFBF or RZFBF based on $\{\mathcal{K}'_g\}$ and
            $\{\mathbf{B}_g\}$;
            Compute the gain
            $\kappa(k, g) = \max\left\{0, \mathcal{C}(\{\mathcal{K}'_g\}, \{\mathbf{B}_g\}) - \mathcal{C}(\{\mathcal{K}_g\}, \{\mathbf{B}_g\})\right\}$;
        end
    end
    Obtain $(k^*, g_{k^*}) = \arg\max_{k \in \mathcal{U}} \kappa(k, g)$ ;
    if $(k^*, g_{k^*}) \neq \emptyset$ then
        $\mathcal{U} \leftarrow \mathcal{U} \backslash k^*$ ;
        $\mathcal{K}_{g_{k^*}} \leftarrow \mathcal{K}_{g_{k^*}} \cup \{k^*\}$ ;
    end
end
```

Algorithm 1. Greedy Algorithm for Dynamic User Selection and Beamforming.

We develop a user grouping method considering group load balancing and user proportional fairness. The problem can be formulated as:

$$\begin{aligned} \max_{\{x_{kg}\}} \quad & \mathcal{J} = \sum_{k=1}^{K} \sum_{g=1}^{G} x_{kg} \log\left(\frac{\bar{\eta}_{gk}}{\sum_i x_{ig}}\right) \\ s.t. \quad & \sum_g x_{kg} = 1, \quad \forall k \in \{1, 2, \cdots, K\} \\ & x_{kg} \in \{0, 1\}, \quad \forall k, g, \end{aligned}$$

g is *1* if user k is connected to group g, and *0* otherwise; $\bar{\eta}_{k_g}$ is the rate for user k in group g. By adopting a two-tier dual decomposition approach, we could obtain the optimal solution to the above problem.

3. Performance Evaluation

Simulations are performed to evaluate the proposed schemes. We fix $M = 100$ and group number G=6.

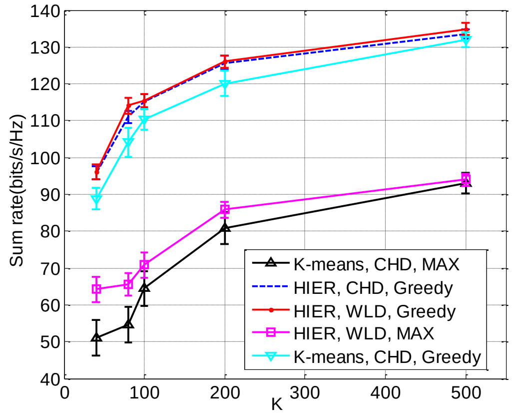

Fig. 1 System sum rate versus the number of users.

We can see from Fig. 1 that all our proposed schemes outperform the scheme in [7]. In particular, hierarchical clustering greedy user selection with weighted likelihood has the highest system throughput.

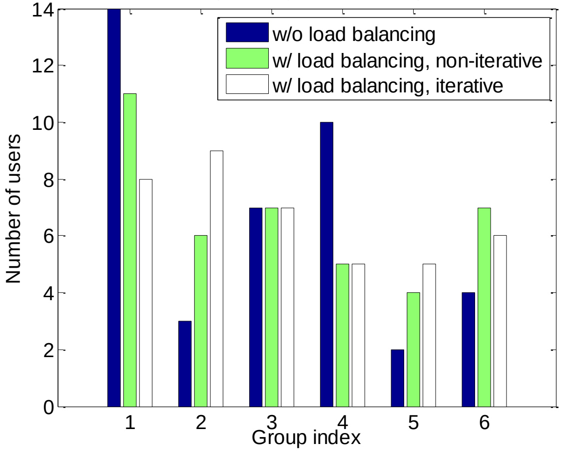

Fig. 2 Group sizes for user grouping with joint group load balancing and precoding design when $K = 40$.

We can see from Fig. 2 that the maximum difference of the scheme without considering load balancing is 12, while this number is 4 in our proposed iterative scheme. Thus the loading of users is much more balanced in our proposed scheme.

4. Conclusions

In this letter, we have studied the user grouping and scheduling problems based on a two-stage precoding framework for FDD massive MIMO systems. We have proposed weighted likelihood similarity measure and hierarchical clustering for user grouping. We have also proposed a dynamic user scheduling scheme and a user grouping algorithm to achieve load balancing and user fairness for FDD massive MIMO systems. The efficacy of the proposed schemes has been validated with analysis and simulations.

ACKNOWLEDGMENT

This work was supported in part by the U.S. National Science Foundation under Grants CNS-1247955 and CNS-1320664.

References

[1] E. Larsson *et al*, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 185--195, Feb. 2014.

[2] F. Rusek *et al*., “Scaling up MIMO: Opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40--60, Jan. 2013.

[3] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640--2651, Aug. 2011.

[4] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, “Inter-cell interference in noncooperative TDD large scale antenna systems,” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192--201, Feb. 2013.

[5] J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, “Making smart use of excess antennas: Massive MIMO, small cells, and TDD,” *Bell Labs Tech. J.*, vol. 18, no. 2, pp. 5--21, Sep. 2013.

[6] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, “Joint spatial division and multiplexing: The large-scale array regime,” *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441--6463, Oct. 2013.

[7] A. Adhikary and G. Caire, “Joint spatial division and multiplexing: Opportunistic beamforming and user grouping,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, iss. 5, pp. 876--890, Mar. 2014.

[8] Y. Xu, G. Yue, N. Prasad, S. Rangarajan, and S. Mao, “User grouping and scheduling for large scale MIMO systems with two-stage precoding,” in *Proc. IEEE ICC*, Sydney, Australia, Jun. 2014, pp. 5208--5213.

[9] Y. X., G. Yue, and S. Mao, “User grouping for Massive MIMO in FDD systems: New design methods and analysis,” *IEEE Access Journal, Special Section on 5G Wireless Technologies: Perspectives of the Next Generation Mobile Communications and Networking,* vol.2, pp.947-959, Aug. 2014.

YI XU (S’11) received the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2010, and the B.S. degree in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2007. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His research interests include optimization, game theory, MIMO, OFDM, IDMA, and cognitive radio networks.

GUOSEN YUE (S'99-M'04-SM'09) received the B.S. degree in physics and the M.S. degree in electrical engineering from Nanjing University, Nanjing, China in 1994 and 1997, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, in 2004. He was a senior research staff with the Mobile Communications and Networking Research Department, NEC

Laboratories America, Princeton, New Jersey, conducting research for broadband wireless systems and mobile networks. His research interests are in the general areas of wireless communications and signal processing. In August 2013, he joined Broadcom Corporation as a system design scientist. Dr. Yue serves as an Associate Editor for the IEEE Transactions on Wireless Communications. He has served as the associate Editor for Research Letters in Communications, the Guest Editors for EURASIP Journal of Wireless Communication and Networking special issue on interference management, ELSEVIER PHYCOM special issue on signal processing and coding. He served as the Symposium Co-chair for IEEE ICC 2010, the Track Co-chair for IEEE ICCCN 2008, the steering committee member for IEEE RWS 2009. He is a senior member of the IEEE.

SHIWEN MAO (S'99-M'04-SM'09) received the Ph.D. degree in electrical and computer engineering from Polytechnic University, Brooklyn, NY, USA. Currently, he is the McWane Associate Professor with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA. His current research interests include cross-layer optimization of wireless networks and multimedia communications, with current focus on cognitive radios, femtocells, 60 GHz mmWave networks, free space optical networks, and smart grid. He is on the Editorial Board of the IEEE Transactions on Wireless Communications, IEEE Internet of Things Journal, IEEE Communications Surveys and Tutorials, and several other journals. He received the 2013 IEEE ComSoc MMTC Outstanding Leadership Award and the NSF CAREER Award in 2010. He is a co-recipient of the IEEE ICC 2013 Best Paper Award and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems.